









28  
414  
3649-  
94

The Parsimonious Property of Cut Covering  
Problems and its Applications

Dimitris Bertsimas  
and  
Chungpiaw Teo

WP# - 3649-94 MSA      January, 1994



The Parsimonious Property of Cut Covering  
Problems and its Applications

Dimitris Bertsimas  
and  
Chungpiaw Teo

WP# - 3649-94 MSA      January, 1994

M.I.T. LIBRARIES

FEB 11 1994

RECEIVED



# The parsimonious property of cut covering problems and its applications

Dimitris Bertsimas \*      Chungpiaw Teo †

January 1994

## Abstract

We consider the analysis of linear programming relaxations of a large class of combinatorial problems that can be formulated as problems of covering cuts, including the Steiner tree, the traveling salesman, the vehicle routing, the matching, the  $T$ -join and the survivable network design problem, to name a few. We prove that all of the problems in the class satisfy a deep structural property, the *parsimonious* property, generalizing earlier work by Goemans and Bertsimas [3]. We identify two set of conditions for the parsimonious property to hold and offer two proof techniques based on combinatorial and algebraic arguments. We examine several consequences of the parsimonious property in proving monotonicity properties of LP relaxations, giving genuinely simple proofs of integrality of polyhedra in this class, offering a unifying understanding of results in disjoint path problems and in the approximability of problems in the class. We also propose a new proof method that utilizes the parsimonious property for establishing worst case bounds between the gap of the IP and LP values. Our analysis unifies and extends a large set of results in combinatorial optimization.

---

\*Dimitris Bertsimas, Sloan School of Management and Operations Research Center, MIT, Cambridge, MA 02139. Research partially supported by a Presidential Young Investigator Award DDM-9158118 with matching funds from Draper Laboratory.

†Chungpiaw Teo, Operations Research Center, MIT, Cambridge, MA 02139.

# 1 Introduction

We consider the following class of problems defined on a graph  $G = (V, E)$  and described by the following integer programming formulation

$$\begin{aligned}
 IZ_f(D) = \quad & \text{minimize} && \sum_{e \in E} c_e x_e \\
 & \text{subject to} && \sum_{e \in \delta(i)} x_e = f(i), \quad i \in D \subseteq V \\
 & && \sum_{e \in \delta(S)} x_e \geq f(S), \quad S \subset V \\
 & && x_e \in Z_+.
 \end{aligned}$$

where  $f : 2^V \rightarrow Z_+$  is a given set function,  $\delta(S) = \{e = (i, j) \in E \mid i \in S, j \in V \setminus S\}$ . By selecting different set functions  $f(S)$  and different sets  $D$  we can model a large class of combinatorial problems, including the Steiner tree, the traveling salesman, the vehicle routing, the matching,  $T$ -join and survivable network design problem, to name a few.

Let  $IP_f(D)$  be the underlying feasible space. We denote the LP relaxation as  $P_f(D)$ , in which we replace constraints  $x_e \in Z_+$  with  $x_e \geq 0$ . We denote the value of the LP relaxation as  $Z_f(D)$ . Goemans and Bertsimas [3] studied the survivable network design problem, in which the objective is to design a network at minimal cost that satisfies connectivity requirements (for each pair  $(i, j)$  of nodes in  $V$ , the solution should contain at least  $r_{ij}$  edge disjoint paths) and considered an integer programming formulation of the type  $IP_f(D)$  with  $f(S) = \max_{(i,j) \in \delta(S)} r_{ij}$ ,  $D = \emptyset$ . They showed the following property, which they call *the parsimonious property*, which in our notation can be stated:

**Theorem 1 [Goemans and Bertsimas [3]]** *If the costs  $c_e$  satisfy the triangle inequality ( $c_{ij} \leq c_{ik} + c_{kj}$  for all  $i, j, k \in V$ ), then for the survivable network design problem ( $f(S) = \max_{(i,j) \in \delta(S)} r_{ij}$ )*

$$Z_f(D) = Z_f(\emptyset).$$

In other words, the degree constraints are unnecessary for the LP relaxation in the survivable network design problem. They further examine several sometimes surprising structural and algorithmic properties of the LP relaxation, and examine the worst case behavior of  $\frac{IZ_f(\emptyset)}{Z_f(\emptyset)}$  for the survivable

network design problem. Goemans and Williamson [4], Williamson et.al. [16] and Goemans et. al. [5] show interesting worst case bounds on the ratio  $\frac{1Z_f(\emptyset)}{Z_f(\emptyset)}$ .

Our goal in this paper is to understand the class of problems for which the parsimonious property holds and examine several implications of the parsimonious property. In this way we shed new light to a large collection of results in discrete optimization and graph theory, understand their common origin and generalize them in interesting ways. In particular our contributions in this paper are:

1. We continue the program started in [3] by identifying a set of conditions on the set function  $f(S)$ , for which the parsimonious property holds. In this way we prove that a large collection of classical combinatorial problems satisfy it including the matching problem, the  $T$ -join problem, a relaxation of the vehicle routing problem, some disjoint path problems, some  $b$ -matching problems, etc. In particular all problems considered in Goemans and Williamson [4] satisfy it. We also find that if the set function  $f(S)$  does not satisfy this set of conditions, the property does not hold. We offer two proofs of the property: a combinatorial proof based on splitting techniques originated in Lovász [8] and used in Goemans and Bertsimas [3] and an algebraic proof based on linear programming duality. Goemans [6] has also independently developed this generalization using the techniques in [3]. The duality proof reveals a further generalization of the parsimonious property to integer programming programs as well (the dual parsimonious property).
2. We use the parsimonious property to prove interesting monotonicity properties of the LP relaxation for problems in this class.
3. We use the parsimonious property to give genuinely simple proofs of the integrality of some polyhedra  $P_f(D)$ : the  $T$ -join problem, special cases of the Steiner tree problem including the shortest path problem and the shortest path tree problem.
4. We further extend the parsimonious property under more general conditions and examine its implications in the disjoint path problem. We find that this extension is the source for several results in this area and provides a unifying framework to understand these results.

5. We offer a new proof technique that utilizes the parsimonious property to find bounds on the ratio  $\frac{IZ_f(\emptyset)}{Z_f(S)}$ . Our proof technique leads to a new approximation algorithm for this class of problems that compared with the algorithm proposed by Goemans and Williamson [4] is simpler to implement as it does not use reverse deletions, but only shortest path computations.
6. We use the parsimonious property to prove new approximation bounds for Problem  $IP_f(D)$  with  $D \neq \emptyset$ .

The paper is structured as follows. In Section 2 we introduce the properties of the set function  $f(S)$  that imply the parsimonious property and examine classical combinatorial problems that can be modelled in this way. In Section 3 we prove the parsimonious property as well as the dual integral parsimonious property. In Section 4 we derive interesting monotonicity properties of this class of problems as consequences of the parsimonious property. In Section 5 we further extend the parsimonious property and apply it to the analysis of the disjoint path problem. In Section 6 we examine applications of the parsimonious property to the integrality of certain polyhedra  $P_f(D)$ . In Section 7 we introduce a new proof technique to bound the ratio  $\frac{IZ_f(\emptyset)}{Z_f(\emptyset)}$ . This proof technique gives rise to a new approximation algorithm for the problems considered in Goemans and Williamson [4]. We further examine applications of the parsimonious property to the approximability of Problem  $IP_f(D)$ .

## 2 Parsimonious set functions

In their study of the approximability of problems in the class  $IP_f(\emptyset)$ , Goemans and Williamson [4] (for the case that  $f(S)$  takes values in  $\{0, 1\}$ ) and Williamson et. al. in [16] (for the case that  $f(S)$  takes values in  $Z_+$ ) introduce the following set of conditions for the set function  $f(S)$ .

**Conditions A (proper set functions):**

1.  $f(\emptyset) = 0$ .
2. **Symmetry:**  $f(S) = f(V \setminus S)$  for all  $S \subseteq V$ .
3. **Properness:** If  $A \cap B = \emptyset$ , then  $f(A \cup B) \leq \max\{f(A), f(B)\}$ .

We next introduce the following set of conditions:

**Conditions B (parsimonious set functions):**

1.  $f(\emptyset) = 0$ .
2. **Symmetry:**  $f(S) = f(V \setminus S)$  for all  $S \subseteq V$ .
3. **Node Subadditivity (NS):** If  $A \cap \{x\} = \emptyset$ , then  $f(A \cup \{x\}) \leq f(A) + f(\{x\})$ .
4. **Quasi-supermodularity (QS):** For all  $S, T \subseteq V, S \cap T \neq \emptyset$

Either

$$f(S) + f(T) \leq f(S \cup T) + f(S \cap T)$$

or

$$f(S) + f(T) \leq f(S \setminus T) + f(S \cap T).$$

We also introduce the general subadditivity condition:

**Subadditivity:** If  $A \cap B = \emptyset$ , then  $f(A \cup B) \leq f(A) + f(B)$ .

The QS property was also introduced in the recent paper of Goemans et. al. [5], who used the term weakly supermodular.

Finally we introduce a third set of conditions on the set function  $f$ :

**Conditions C (weakly parsimonious set functions):**

1.  $f(\emptyset) = 0$ .
2. **Symmetry:**  $f(S) = f(V \setminus S)$  for all  $S \subseteq V$ .
3. **Weak Subadditivity (WS):** If  $A \cap \{x\} = \emptyset$ , then  $f(A \cup \{x\}) \leq f(A)$ . We will then say that  $x$  is a weakly Steiner vertex.
4. **2-Quasi-supermodularity (2-QS):** For every three mutually crossing sets (two sets  $A, B$  are crossing if  $A \setminus B, B \setminus A, A \cap B$  are nonempty) at least two of them satisfy the QS property.

Compared with Conditions B (parsimonious set functions), weak subadditivity is stronger than node subadditivity, while the 2-QS property is a relaxation of the QS property. In other words, there are set functions  $f$  satisfying one of conditions B or C but not the other.

We next show that Conditions *B* are more general than Conditions *A*.

**Proposition 1** [Goemans et. al. [5]] *Let  $f$  be a symmetric, set function with  $f(\emptyset) = 0$ . Then, if  $f$  is proper, it is quasisubmodular and node subadditive.*

**Proof :** If  $f$  is a proper function, then clearly  $f$  is node subadditive. Among the terms  $f(S \cap T)$ ,  $f(S \cup T)$ ,  $f(S \setminus T)$ ,  $f(T \setminus S)$ , say  $f(S \cap T)$  attains the minimum. By properness,  $f(S) \leq \max(f(S \cap T), f(S \setminus T)) = f(S \setminus T)$ , and  $f(T) \leq \max(f(S \cap T), f(T \setminus S)) = f(T \setminus S)$ , and so  $f(S) + f(T) \leq f(S \setminus T) + f(T \setminus S)$ . The other cases follow similarly from symmetry of  $f$ .  $\square$

In Figure 1 (we consider symmetric set functions) we draw the relations of the various conditions we considered. As we show in the next sections, the parsimonious property holds for set functions  $f$  satisfying either conditions *A*, or *B*, or *C*.

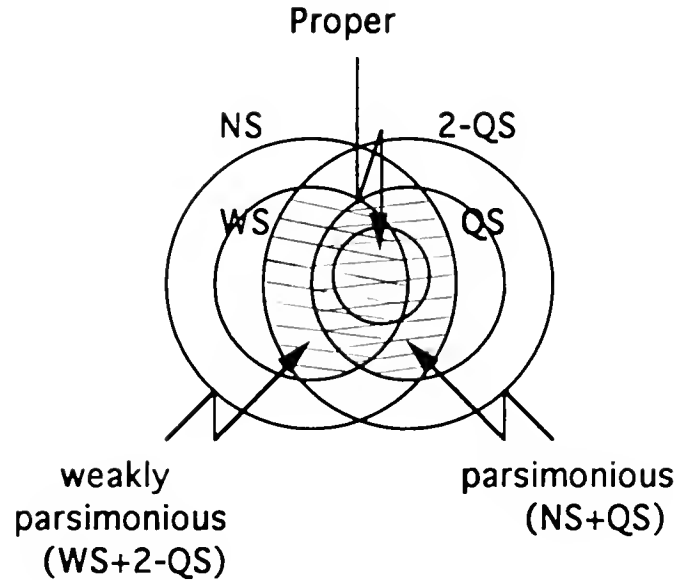


Figure 1: Relations of parsimonious, weakly parsimonious and proper set functions.

**Remarks:**

1. The following simple observation is usually useful in checking whether a given function has the QS property. We select any node  $v$  and check the QS property only for those sets  $S$  and  $T$  containing  $v$ . By symmetry of  $f$ , we can extend the QS property to all  $S$  and  $T$ .

2. Conditions  $B$  are strictly more general, i.e., there are set functions which are parsimonious but not proper.
3. The symmetry conditions are without loss of generality. If the function  $f$  is not symmetric, we can redefine the following symmetric set function:  $\hat{f}(S) = \hat{f}(V \setminus S) = \max[f(S), f(V \setminus S)]$ . Notice that  $I Z_{\hat{f}}(\emptyset) = I Z_f(\emptyset)$  and  $Z_{\hat{f}}(\emptyset) = Z_f(\emptyset)$ , since the optimal solution of  $IP_f(\emptyset)$  is feasible in  $IP_{\hat{f}}(\emptyset)$  and vice versa.

## 2.1 Examples of problems

In Table 1 below we review several classical combinatorial problems formulated using the cutset formulation  $IP_f(\emptyset)$  for  $f$  satisfying both Conditions  $A$  and  $B$ . In Section 5 we show how Conditions  $C$  naturally arise in the study of the disjoint path problem.

Problem	$f(S)$	Conditions
Spanning tree	1 for all $S \neq \emptyset, V$	$A, B$
Steiner tree	1 if $S \cap T \neq \emptyset, T$	$A, B$
Shortest path	1, $ S \cap \{s, t\}  = 1$	$A, B$
Generalized Steiner tree	1 if $S \cap T_i \neq \emptyset, T_i, i = 1, \dots, k$	$A, B$
Nonbipartite matching	1 if $ S  = 2k + 1$	$A, B$
$T$ -join	1 if $ S \cap T  = 2k + 1$	$A, B$
Network survivability	$\max_{(i,j) \in \delta(S)} r_{i,j}, r_e \geq 0$	$A, B$
$k$ -connected graph	$k$ for all $S \neq \emptyset, V$	$A, B$
Tree partitioning	1, if $ S  \leq k,  S  \geq n - k$	$A, B$
Point-to-point connection	1, if $ S \cap C  \neq  S \cap D $	$A, B$

Table 1: Problems formulated as  $IP_f(\emptyset)$  satisfying Conditions  $A$  (and therefore  $B$ ).

We next describe problems that are parsimonious but not proper.

### The $b$ -matching problem

Given numbers  $b(i)$  such that  $\sum_{i \in V} b(i) = 2r$ , the problem can be modeled in the form of  $IP_f(D)$

with  $D = V$  and

$$f(S) = \begin{cases} 1 & |S| \geq 2, \sum_{i \in S} b(i) = 2k+1, \\ b(i) & S = \{i\}, V \setminus \{i\} \end{cases}$$

Notice that the function  $f$  is not proper, because for  $A, B$  disjoint whose union is  $V \setminus \{i\}$  the definition is violated. However, the function  $f(S)$  is QS. While  $f$  is not subadditive for general sets  $A, B$ , it is node subadditive if  $b(i) \in \{a, a+1\}$ . While in general the  $b$ -matching problem does not satisfy the parsimonious property, it does satisfy it if  $b(i) \in \{a, a+1\}$ .

### The capacitated tree problem

Given a graph  $G = (V \cup \{0\}, E)$ , demands  $d_i$ ,  $i \in V$ , a depot 0, costs  $c_e$ ,  $e \in E$  and we would like to design a tree of minimum cost such that each subtree from the depot has demand at most  $Q$ . The capacitated tree problem is a popular relaxation of the vehicle routing problem. A valid cutset formulation of the capacitated tree problem is of the type  $IP_f(\emptyset)$  with

$$f(S) = \begin{cases} 2 \frac{\sum_{i \in S} d_i}{Q} & S \subset V, \\ 2 \frac{\sum_{i \in V \setminus S} d_i}{Q} & 0 \in S \end{cases}$$

It is obvious that  $f(S)$  does not satisfy Conditions A, since it is not proper but it satisfies Conditions B: It is clearly symmetric and subadditive as we show below: For  $S, T \subset V$  such that  $S \cap T = \emptyset$

$$f(S \cup T) = 2 \frac{\sum_{i \in S \cup T} d_i}{Q} = f(S) + f(T),$$

$$f((S \cap \{0\}) \cup T) = 2 \frac{\sum_{i \in V \setminus (S \cup T)} d_i}{Q} \leq f(S \cap \{0\}) \leq f(S \cap \{0\}) + f(T).$$

It is also QS, since for  $S, T \subset V$  containing 0,

$$f(S \cup T) + f(S \cap T) = f(S) + f(T).$$

### The traveling salesman and vehicle routing problem

The traveling salesman problem can be modeled as  $IP_f(D)$  with  $D = V$  and  $f(S) = 2$  for all  $S \subset V$ . Interestingly, the vehicle routing problem can be modeled in our framework as follows. Given a graph  $G = (V \cup \{0\}, E)$ , demands  $d_i$ ,  $i \in V$ , a depot 0, costs  $c_e$ ,  $e \in E$  and vehicles of capacity  $Q$  we want to find tours of the vehicles from the depot of minimum cost, such that



the demand in each tour does not exceed capacity. Notice that the capacitated tree problem is a relaxation of the vehicle routing problem. We can strengthen the formulation if we write for example  $f(S) = 2\lceil \frac{\sum_{i \in S} d_i}{Q} \rceil$ . While this set function is subadditive, it is not QS.

### 3 The parsimonious property

The cut-set formulation introduced in the previous section captures many of the classical optimization problems studied in the literature. It is thus interesting and indeed surprising that the parsimonious property holds for the LP relaxations of these problems. In the remainder of this section, we prove the parsimonious property in two ways:

1. The primal proof is an extension of edge splitting techniques introduced in Lovász [8] and used in [3] to prove the parsimonious property for the survivable network design problem;
2. The dual proof uses linear programming duality and extends an observation of Frank [2] for the matching problem to the general class of problems  $P_f(\emptyset)$ .

#### 3.1 A primal proof of the parsimonious property

Let  $x$  be a feasible solution in  $P_f(\emptyset)$  with  $x(v, u), x(v, w) > 0$ , where  $u, v$  and  $w$  are vertices in  $G$ . We split  $v$  at  $\{u, w\}$  by some  $\Delta > 0$  in the following way :

$$\begin{aligned} x(v, u) &\leftarrow x(v, u) - \Delta \\ x(v, w) &\leftarrow x(v, w) - \Delta \\ x(u, w) &\leftarrow x(u, w) + \Delta. \end{aligned}$$

The splitting operation will preserve feasibility of  $x$  unless there exists a set  $S \subset V$  such that  $v \in S, u, w \notin S$  and  $x(\delta(S)) = f(S)$ . We call such a set  $S$  a tight set. We denote by  $\overline{S}$  the complement of  $S$ . We also use the notation  $x(\delta(A, B)) = \sum_{e=\{i,j\}, i \in A, j \in B} x_e$  and  $x(\delta(S)) = x(\delta(S, \overline{S}))$ .

We need a preliminary lemma regarding properties of tight sets.

**Lemma 2** *If  $S, T$  are tight sets, and  $f$  is QS, then*

*either (i)  $S \setminus T, T \setminus S$  are tight,  $x(\delta(S \cap T, \overline{S \cup T})) = 0$*

*or (ii)  $S \cap T, S \cup T$  are tight,  $x(\delta(S \setminus T, T \setminus S)) = 0$ .*

**Proof :** Since  $f$  is QS, we first consider the case  $f(T \setminus S) + f(S \setminus T) \geq f(S) + f(T)$ . In this instance,

$$\begin{aligned} f(T \setminus S) + f(S \setminus T) &\leq x(\delta(T \setminus S)) + x(\delta(S \setminus T)) \\ &= x(\delta(S)) + x(\delta(T)) + 2x(\delta(S \cap T, \overline{S \cup T})) \\ &= f(S) + f(T) + 2x(\delta(S \cap T, \overline{S \cup T})). \end{aligned}$$

Hence  $x(\delta(S \setminus T)) = f(S \setminus T)$ ,  $x(\delta(T \setminus S)) = f(T \setminus S)$  and  $x(\delta(S \cap T, \overline{S \cup T})) = 0$ , i.e., condition (i) holds. On the other hand, the case  $f(S \cup T) + f(S \cap T) \geq f(S) + f(T)$  gives rise to condition (ii), using an identical argument.  $\square$

We can now prove the central result of this section.

**Theorem 3 (Parsimonious Property)** *If the cost function  $c$  satisfies the triangle inequality, and  $f$  is a parsimonious set function, then*

$$Z_f(D) = Z_f(\emptyset), \quad \text{for all } D.$$

**Primal proof:** Let  $x$  be an optimal solution in  $P_f(\emptyset)$ , with  $\sum_{e \in E} x_e$  minimal. Suppose there is a  $v$  in  $D$  that has  $x(\delta(v)) > f(v)$ ; Let  $u$  be such that  $x(v, u) > 0$ . Let  $S$  be a minimal tight set containing  $v$  but not  $u$ . (If such set does not exist, then we can decrease  $x(v, u)$  by some positive  $\Delta$ , while maintaining feasibility. This contradicts the minimality of  $x$ .) If  $T$  is another tight set containing  $v$  but not  $u$ , then  $x(\delta(S \cap T, \overline{S \cup T})) \geq x(v, u) > 0$ . Since  $f$  is QS, condition (ii) of Lemma 2 holds and so  $S \cap T$  is a tight set. By the minimality of  $S$ ,  $S$  is contained in  $T$ . So there is a unique minimal tight set containing  $v$  but not  $u$ .

In addition, there exists a  $w$  in  $S$  with  $x(v, w) > 0$ , else  $f(S) = x(\delta(S)) = x(\delta(v)) + x(\delta(S \setminus \{v\})) > f(v) + f(S \setminus \{v\})$ , violating the node subadditivity of  $f$ . We can then split  $v$  at  $\{u, w\}$  by some positive  $\Delta$ , where

$$\Delta \leq \frac{1}{2} \min\{x(\delta(S)) - f(S) : v \in S, u, w \notin S, x(\delta(S)) - f(S) > 0\}.$$

Because of the triangle inequality, this operation yields another feasible optimal solution  $x'$  with  $x'(\delta(v)) < x(\delta(v))$  and  $x'(\delta(t)) = x(\delta(t))$  if  $t \neq v, t \in V$ . This contradicts the minimality of  $x$ .  $\square$

**Remarks:**

1. In the splitting process, if  $x, f$  are both even and integral, we can choose  $\Delta$  to be 1. This corresponds to the classical edge-splitting notion and has played an important role in many connectivity problems. We summarize this discussion in the following corollary.

**Corollary 1** *Let  $G$  be an Eulerian multigraph and  $x_G$  be the incidence vector of  $G$ . Let  $f$  be an even, parsimonious set function. If  $x_G$  is a feasible integral solution to  $IP_f(\emptyset)$  and  $x(\delta(v)) > f(\{v\})$  for some  $v \in V$ , then there exists  $\{u, w\}$  and an edge splitting operation of  $v$  at  $\{u, w\}$ , yielding a new graph  $G'$  and a corresponding incidence vector  $x_{G'}$  that is a feasible integral solution to  $IP_f(\emptyset)$ .*

This corollary generalizes the following result of Lovász [8] and a refinement due to Goemans and Bertsimas [3].

**Corollary 2** *Let  $G$  be an Eulerian multigraph,  $r_G(i, j)$  be the maximum number of edge disjoint paths from  $i$  to  $j$  and  $v$  be a vertex of  $G$ . Then there exists  $\{u, w\}$  and an edge splitting operation of  $v$  at  $\{u, w\}$  (obtaining a new graph  $G'$ ) such that*

- $r_{G'}(i, j) = r_G(i, j)$  if  $i, j \neq v$ .
- $r_{G'}(v, j) = \min(r_G(i, j), \deg_G(v) - 2)$  for  $j \neq v$ .

**Proof :** Set  $f(S) = \max\{r_G(i, j) : e = (i, j) \in \delta(S)\}$ . Note that  $f$  is a parsimonious function. The incidence vector  $x_G$  of graph  $G$  is a feasible integral solution in  $IP_f(\emptyset)$ , with  $x(\delta(v)) = \deg_G(v) > f(v)$ . Since  $G$  is Eulerian,  $f(S)$  and  $x(\delta(S))$  are even for each  $S \subset V$ . Thus there exists an edge splitting operation of  $v$ . The graph  $G'$  obtained in this way is a feasible solution with  $x_{G'}(\delta(S)) \geq f(S)$ . By the max-flow-min-cut theorem, the graph  $G'$  has the required connectivity.  $\square$

2. Notice that for the parsimonious property to hold  $f$  needs to be QS and node subadditive. The full subadditivity is not needed. The  $b$ -matching problem, for example with  $b(v) = k$  or  $k + 1$  is both QS and node subadditive, while if  $b(v) \in \{k, k + 1, k + 2\}$  it is not node subadditive.

Although the condition that the cost function  $c$  satisfies the triangle inequality seems restrictive, we next show that for problems of the form  $IP_f(\emptyset)$ ,  $P_f(\emptyset)$ , i.e., with no degree constraints, we can ensure that this condition is met by the following transformation. Let  $c'(u, v)$  denote the shortest path between  $u$  and  $v$  with  $c$  as the length function. Clearly  $c'$  satisfies the triangle inequality. Let  $I Z'_f(\emptyset)$  and  $Z'_f(\emptyset)$  denote the respective solution value with  $c'$  as the objective function.

**Theorem 4**

$$Z'_f(\emptyset) = Z_f(\emptyset); \quad I Z'_f(\emptyset) = I Z_f(\emptyset).$$

**Proof :** Let  $x$  be an optimal solution to  $P'_f(\emptyset)$  (or  $IP'_f(\emptyset)$ ). Consider an edge  $e = (u, v)$  such that  $c'(e) < c(e)$ . Let  $P$  be a shortest path (with respect to  $c$ ) linking  $u$  and  $v$ . Then  $P \neq \{e\}$ , and  $c'(g) = c(g)$  for each edge  $g$  on  $P$ . If  $x(e) > 0$ , we can reroute the flow through the path  $P$ , resulting in another optimal solution (since  $c'(e) = c'(P)$ ). Repeating this procedure, we have  $x(e) > 0$  only when  $c'(e) = c(e)$ , thus proving the theorem.  $\square$

### 3.2 A dual proof of the parsimonious property

The dual of  $P_f(D)$  is as follows :

$$\begin{aligned} DZ_f(D) = \text{maximize} \quad & \sum_{S \subset V} y(S) f(S) \\ \text{subject to} \quad & \sum_{S: e \in \delta(S)} y(S) \leq c(e), e \in E \\ & y(S) \geq 0, S \subset V, S \notin D. \end{aligned}$$

Let  $DP_f(D)$  be the dual polyhedron and  $DZ_f(D)$  denote the optimal objective value. To prove the parsimonious property using a dual argument, we only need to show that among all dual optimal solutions to  $DP_f(D)$ , we can always choose one with  $y(v) \geq 0$  for all  $v \in D$ . This solution is then feasible to  $DP_f(\emptyset)$ . Let  $\mathcal{T}$  be a collection of sets (subsets of  $V$ ). We call this family of sets **laminar** if for all  $A, B \in \mathcal{T}$  either  $A \cap B = \emptyset$ , or  $A \subset B$ , or  $B \subset A$ .

**Theorem 5** *If the cost function  $c$  satisfies the triangle inequality, and  $f$  is a parsimonious set function, then*

$$Z_f(D) = Z_f(\emptyset), \quad \text{for all } D.$$

**Dual proof:** Let  $y$  be a dual optimal solution in  $DP_f(D)$ . By the QS property, we may assume that the set  $\mathcal{F} := \{S : y(S) > 0\}$  is laminar, since we can always replace two intersecting sets  $S$  and  $T$  by  $S \setminus T, T \setminus S$  or  $S \cap T, S \cup T$ . Suppose there exists a  $v \in D$  such that  $y(v) < 0$ . For all  $A \in \mathcal{F}$  containing  $v$ , we replace  $A$  by  $V \setminus A$ , i.e., we set

$$y(A) \leftarrow 0, \quad y(V \setminus A) \leftarrow y(V \setminus A) + y(A).$$

In this way we obtain another dual optimal solution with no member of  $\mathcal{F}$  containing  $v$ . Note that  $\mathcal{F}$  is still laminar.

Let  $p(e) = \sum_{S:e \in \delta(S)} y(S)$ . By dual feasibility,  $p(e) \leq c(e)$ . We may assume that there is a  $u \in V$  such that  $p(u, v) = c(u, v)$ , since we can increase  $y(v)$  otherwise. Let  $A$  be a maximal member of  $\mathcal{F}$  containing  $u$ . Let  $\Delta = \min(-y(v), y(A))$ . We modify the dual solution as follows :

$$\begin{aligned} y(v) &\leftarrow y(v) + \Delta \\ y(\overline{A \cup \{v\}}) &\leftarrow y(\overline{A \cup \{v\}}) + \Delta \\ y(A) &\leftarrow y(A) - \Delta. \end{aligned}$$

To check for feasibility of this modified solution, we only need to consider edges of the form  $(v, w)$  where  $w$  is not in  $A$ . Note that by the construction of  $\mathcal{F}$ ,  $p(u, w) = p(v, w) + p(u, v) - 2y(v)$ . Hence  $c(v, w) \geq c(u, w) - c(u, v) = c(u, w) - p(u, v) \geq p(u, w) - p(u, v) = p(v, w) - 2y(v) \geq p(v, w) + 2\Delta$ . Thus the modified solution is dual feasible. By repeating this procedure, we can construct a dual optimal solution with  $y(v) \geq 0$  for all  $v$  in  $D$ .  $\square$

Notice that if  $y$  in the above proof takes only integral values, then  $\Delta$  can be chosen to be integral. This yields an integral analogue of the parsimonious property in a dual sense.

Let  $DIZ_f(D)$  denote the optimal objective value over  $DP_f(D)$  with integrality constraints on  $y(S)$ .

**Theorem 6 (Dual Integral Parsimonious Property)** *If  $f$  is parsimonious, and  $c$  satisfies the triangle inequality, then  $DIZ_f(\emptyset) = DIZ_f(D)$ .*

### 3.3 On the minimality of conditions for the parsimonious property

We remark in this subsection that the parsimonious property does not hold if we relax either the QS or the node subadditivity property.

Consider the set function  $f$  on 3 nodes as follows:  $f(v_1) = f(\{v_2, v_3\}) = 1$ ,  $f(S) = 0$  otherwise. Then clearly  $f$  is QS, but it is not node subadditive. In this case the parsimonious property does not hold, as the polyhedron  $P_f(V)$  is empty.

On the other hand, subadditivity alone does not guarantee the parsimonious property. Define  $f$  on 4 nodes as follows:  $f(S) = 1$  if  $|S| = 1$  or 3,  $f(S) = 2$  otherwise. Then  $f$  is clearly subadditive and symmetric. In this instance,  $P_f(V)$  is again empty, since if  $x(v_i, v_j) > 0$  and  $x(v_i) = x(v_j) = 1$ , then  $x(\delta(\{v_i, v_j\})) < 2$ .

## 4 Monotonicity properties

Let  $f$  be a parsimonious set function defined on  $V$ , and let  $W$  be a set disjoint from  $V$ . Let  $f_W(S) = f(S \cap V)$  for  $S \subset V \cup W$ . We call  $f_W$  an extension of  $f$  to  $V \cup W$ . It is easy to check that  $f_W$  is again a parsimonious set function. Note that  $f_W(S) = 0$  if  $S \subset W$ . For this reason, we call  $W$  the set of Steiner vertices.

For instance, when  $f(S) = 1$  for all  $S \subset V$ , the extension  $f_W$  corresponds to the Steiner tree problem, with  $W$  the set of Steiner vertices. When  $f(S) = 1$  for all odd  $S$  in  $V$ ,  $f_W$  is the  $V$ -join function. Goemans and Bertsimas [3] for the survivable network design problem and independently Shmoys and Williamson [15] for the Held and Karp bound proved the following monotonicity result. We extend this result to the class of parsimonious functions.

**Theorem 7** *Let  $f, g$  be parsimonious functions defined on  $V$  and  $V \cup W$  respectively and  $f_W$  defined as above. Suppose  $f_W(S) \leq g(S)$  for all  $S \subset V \cup W$ . If the cost function  $c$  (defined on  $V \cup W$ ) satisfies the triangle inequality, then*

$$Z_f(V) \leq Z_g(V \cup W).$$

**Proof :**

$$\begin{aligned} Z_f(V) &= Z_{f_W}(V \cup W) \\ &= Z_{f_W}(\emptyset) \quad (\text{parsimonious property}) \\ &\leq Z_g(\emptyset) \quad (g \text{ dominates } f_W) \\ &= Z_g(V \cup W) \quad (\text{parsimonious property}). \end{aligned}$$

□

Notice that the monotonicity result does not hold for  $IZ_f(D)$  in general. This is due in part to the fact that the parsimonious property does not hold for integral solutions. On the other hand, since the parsimonious property holds for dual integral solutions, we show next that the following monotonicity result.

**Theorem 8** *Let  $f, f_W, g$  be defined as above and  $f_W \leq g$ . If  $c$  (defined on  $V \cup W$ ) satisfies the triangle inequality, then*

$$DIZ_f(\emptyset) = DIZ_{f_W}(\emptyset) \leq DIZ_g(\emptyset).$$

**Proof :** Clearly  $DIZ_{f_W}(\emptyset) \leq DIZ_g(\emptyset)$ , since  $f_W \leq g$ . We show next that  $DIZ_f(\emptyset) = DIZ_{f_W}(\emptyset)$ . Let  $\{y(S) : S \subset V\}$  be an optimal solution to  $DIP_f(\emptyset)$ . Let  $M$  be a large positive number. We construct a feasible solution to  $DIP_{f_W}(W)$  as follows:

$$\begin{aligned} y'(S) &= y(S) \text{ if } S \subset V, \\ y'(v) &= -M \text{ if } v \in W, \\ y'(S) &= 0 \text{ otherwise.} \end{aligned}$$

Since  $f_W(v) = 0$  for  $v$  in  $W$ ,  $\sum_{S \subset V \cup W} y'(S) f_W(S) = \sum_{S \subset V} y(S) f(S)$ . Therefore,  $DIZ_{f_W}(W) \geq DIZ_f(\emptyset)$ . From the dual parsimonious property,  $DIZ_{f_W}(\emptyset) = DIZ_{f_W}(W) \geq DIZ_f(\emptyset)$ .

On the other hand, if  $\{y(S) : S \subset V \cup W\}$  is optimal for  $DIZ_{f_W}(\emptyset)$ , then by constructing  $y'(S) := \sum_{T: T \cap V = S} y(T)$ , we obtain a feasible solution to  $DIP_f(\emptyset)$ , with the same objective solution. Hence  $DIZ_f(\emptyset) \geq DIZ_{f_W}(\emptyset)$ . We conclude that  $DIZ_f(\emptyset) = DIZ_{f_W}(\emptyset)$ . □

## 5 Weakly parsimonious functions and the disjoint path problem

A natural question is whether node subadditivity and quasisupermodularity (QS) are the most general conditions on the set function  $f$  for the parsimonious property to hold. In this section we show that the parsimonious property still holds for weakly parsimonious set functions (Conditions C introduced in Section 2) and observe that these relaxed conditions provide a unifying understanding of several results on the the disjoint path problem.

We next prove that the parsimonious property holds for weakly parsimonious functions.

**Theorem 9** *Let  $D$  be the set of weakly Steiner vertices. If  $c$  satisfies the triangle inequality, and  $f$  is weakly parsimonious, then*

$$Z_f(D) = Z_f(\emptyset).$$

**Proof :** The proof is similar to the proof of Theorem 3. Let  $v \in D, u \in V$ , and suppose  $x(\delta(v)) > f(v), x(v, u) > 0$ , where  $x$  is an optimal solution in  $P_f(\emptyset)$ . Consider the minimal tight sets that contains  $v$  but not  $u$ . By the 2-QS property, there exist at most 2 such minimal sets, say  $S_1$  and  $S_2$ . Then all tight sets containing  $v$  but not  $u$  must contain one of these two sets.

We show next that there is a  $w$  in  $S_1 \cap S_2$  with  $x(v, w) > 0$ . Assuming the contrary, then

$$\begin{aligned} f(S_i) = x(\delta(S_i)) &= x(\delta(S_i \setminus \{v\})) + x(\delta(\{v\}, \overline{S_i})) - x(\delta(\{v\}, S_i)) \geq \\ & f(S_i \setminus \{v\}) + x(\delta(\{v\}, \overline{S_i})) - x(\delta(\{v\}, S_i)). \end{aligned}$$

From weak subadditive,  $f(S_i \setminus \{v\}) \geq f(S_i)$ ; hence,

$$x(\delta(\{v\}, \overline{S_i})) \leq x(\delta(\{v\}, S_i)).$$

Since we have assumed that  $x(\delta(\{v\}, S_1 \cap S_2)) = 0$ , we rewrite the inequality for  $i = 1, 2$  and obtain

$$x(\delta(\{v\}, S_2 \setminus S_1)) + x(\delta(\{v\}, \overline{S_1 \cup S_2})) \leq x(\delta(\{v\}, S_1 \setminus S_2)),$$

and

$$x(\delta(\{v\}, S_1 \setminus S_2)) + x(\delta(\{v\}, \overline{S_1 \cup S_2})) \leq x(\delta(\{v\}, S_2 \setminus S_1)).$$

Hence  $x(\delta(\{v\}, \overline{S_1 \cup S_2})) \leq 0$ , which is a contradiction since  $x(v, u) > 0$  and  $u \in \overline{S_1 \cup S_2}$ . Therefore, there exists a  $w$  in  $S_1 \cap S_2$  with  $x(v, w) > 0$ . By splitting at  $v$  using  $u, w$ , we obtain another feasible optimal (because of the triangle inequality) solution. By repeating this procedure, we obtain an optimal solution in  $P_f(D)$ , thus proving the theorem.  $\square$

Similar to Corollary 1 the above proof actually yields the following:

**Corollary 3** *Let  $G$  be an Eulerian multigraph and  $x_G$  be the incidence vector of  $G$ . Let  $f$  be an even, 2-QS set function. If  $x_G$  is a feasible integral solution to  $IP_f(\emptyset)$  and  $x(\delta(v)) > f(\{v\})$  for some weakly Steiner vertex  $v \in V$ , then there exists  $\{u, w\}$  and an edge splitting operation of  $v$  at  $\{u, w\}$ , yielding a new Eulerian graph  $G'$  and a corresponding incidence vector  $x_{G'}$  that is a feasible integral solution to  $IP_f(\emptyset)$ .*



As we show next, this corollary provides a unifying way to understand several seemingly unrelated results for the edge-disjoint-path (EDP) problem.

### 5.1 2-QS functions and the disjoint path problem

Given an undirected graph  $G = (V, E)$ , a collection of source-sink pairs  $\{s_1, t_1\}, \dots, \{s_k, t_k\}$ , the EDP problem asks whether there exists a collection of edge disjoint paths in  $G$ , each joining a source to its corresponding sink. Let  $H$  denote the demand graph, with edge set  $\{(s_1, t_1), \dots, (s_k, t_k)\}$ . Let  $x_G(e) = 1$  if  $e \in G$  and let  $x_H(e) = 1$  if  $e \in H$ .

Clearly, a necessary condition for the existence of these paths is the cut-criterion:

$$x_G(\delta(S)) \geq x_H(\delta(S)) \text{ for all } S \subset V.$$

There has been an extensive literature (see for example Frank [2] and Schrijver [13]) that finds conditions on  $G$  and  $H$ , so that the cut-criterion is both necessary and sufficient for the existence of a solution to the EDP problem. Let  $K_n, C_n$  denote respectively the complete graph and the cycle on  $n$  nodes. We also denote the disjoint union of  $m$  copies of  $K_n$  by  $mK_n$ . The following results are known:

**Theorem 10** *If  $G + H$  is Eulerian, and  $H$  is either a double star or a  $K_4$  or a  $C_5$ , then the cut-criterion is necessary and sufficient for the solvability of the EDP problem.*

The case in which  $H$  is a  $2K_2$  was proved by Rothschild and Whinston [12]. The double star case follows easily from their result. The  $K_4$  case was proved by Seymour [14] and Lomonosov [9] independently. The  $C_5$  case is due to Lomonosov [9]. See [2] and [13] for nice proofs and exposition of these results.

At first sight these results appear to be unrelated without a unifying characteristic. We could use the theory developed in this section to identify the unifying characteristic of all the above results contained in Theorem 10. The central reason is that the set function  $x_H(\delta(S))$  in these cases has the 2-QS property. In particular it is easy to prove the following proposition:

**Proposition 2** *The set function  $x_H(\delta(S))$  has the 2-QS property if and only if  $H$  does not contain a  $3K_2$  or disjoint copies of  $K_3$  and  $K_2$ . This in turn holds if and only if  $H$  is a double star or a  $K_4$  or a  $C_5$ .*

In order to see how Proposition 2 can be used to prove Theorem 10, let us rewrite the cut condition as follows:

$$x_{G+H}(\delta(S)) \geq f(S) = 2x_H(\delta(S)).$$

Under the assumptions of Theorem 10 and using Proposition 2,  $x_{G+H}$  corresponds to a Eulerian graph (by assumption), while  $f$  is an even, 2-QS set function. Let  $D = V \setminus \{s_1, t_1, \dots, s_k, t_k\}$  be the nodes in  $G$  that do not belong to a source-sink pair. In our context  $D$  is the set of weakly Steiner vertices. Applying Corollary 3 we can then perform edge-splitting operations on the edges of  $G$  to obtain a new graph  $G'$  that satisfies the cut criterion, but with edges incident only to the sources or sinks. The rest of the proof involves showing that  $G'$  has the set of edge-disjoint-paths joining each source-sink pair. This follows from a tedious case by case analysis which we omit here, as it is unrelated to the theme of the paper. By reversing the edge-splitting operations, we obtain a set of edge-disjoint-paths in  $G$  that meets the cut criterion, thus proving Theorem 10.

## 6 Applications in proofs of integrality of polyhedra

An important direction of research in integer programming is the development of techniques to show integrality of the associated polyhedra for integer programming problems. Perhaps the most common proof technique is algorithmic. Researchers develop an optimal algorithm for a combinatorial optimization problem, which at the same time shows integrality of a proposed formulation for the problem. In this section we show that the parsimonious property leads to *non-algorithmic, genuinely simple* proofs of integrality of some polyhedra  $P_f(D)$ , yielding new simple proofs of some classical results as well as some new results.

A milestone in combinatorial optimization is the proof of integrality (Edmonds [1]) of the perfect matching polyhedron. This result follows directly from the integrality of the  $T$ -join polyhedron, as the perfect matching polyhedron is a face on the  $T$ -join polyhedron. Surprisingly, we can

derive the integrality of  $T$ -join polyhedron from that of the perfect matching polyhedron, using the parsimonious property.

**Theorem 11** *Let  $f(S) = 1$  if  $|S \cap T|$  is odd, 0 otherwise. Then*

$$\text{Conv}(IP_f(\emptyset)) = P_f(\emptyset).$$

**Proof :** Let  $f_T$  be the restriction of  $f$  to  $T$ , defined on  $S \subset T$ . Note that  $IP_{f_T}(T)$  is just the perfect matching polyhedron on  $T$ . We show next that  $IZ_f(\emptyset) = Z_f(\emptyset)$  for all integral cost functions  $c$ . By Theorem 4, we may assume  $c$  satisfies the triangle inequality. The following inequalities are immediate:

$$IZ_f(\emptyset) \leq IZ_{f_T}(\emptyset) \leq IZ_{f_T}(T).$$

From the integrality of the perfect matching problem  $IZ_{f_T}(T) = Z_{f_T}(T)$ . From the parsimonious property  $Z_{f_T}(T) = Z_{f_T}(\emptyset) = Z_f(\emptyset)$ , yielding that

$$IZ_f(\emptyset) \leq Z_f(\emptyset).$$

The reverse inequality holds trivially and so  $IZ_f(\emptyset) = Z_f(\emptyset)$ , which shows integrality of the  $T$ -join polyhedron. □

In the next section we briefly review another (and in our opinion quite powerful) proof technique that proves the integrality of the perfect matching polyhedron directly.

The shortest path polyhedron can be treated as a Steiner-1-connectivity polyhedron on two terminal nodes. Integrality of the polyhedron also follows easily from the parsimonious property. We generalize this result, using the parsimonious property, and show that the cut set formulation for the Steiner-2-Connected polyhedron with at most 5 terminal vertices is integral.

**Theorem 12** *For the Steiner-2-Connected problem on at most 5 terminal nodes,*

$$\text{Conv}(IP_f(\emptyset)) = P_f(\emptyset).$$

**Proof :** It is well known ([11]) that the TSP polyhedron on at most 5 nodes is integral. From the parsimonious property the result follows easily. □

We next consider the multi-commodity flow problem with a single source  $s$ , multiple sinks  $\{t_1, t_2, \dots, t_k\}$  and with no capacity constraints. Let  $D = \{s, t_1, t_2, \dots, t_k\}$ . Algorithmically, the problem reduces to the computation of the corresponding shortest paths between the source and the sinks. Note that if  $D = V$  the problem is the shortest path tree problem. We show that the cut-set formulation for the problem with  $f(S) = \sum_{t_i \in S} 1$  if  $s \notin S$  and  $f(S) = f(\bar{S})$  is integral. This result also follows from Johnson [7].

**Theorem 13** *For the uncapacitated multi-commodity flow problem with a single source and multiple sinks,*

$$\text{Conv}(IP_f(\emptyset)) = P_f(\emptyset).$$

**Proof :** We only need to show that  $IZ_f(D) = Z_f(D)$  when  $c$  satisfies the triangle inequality. Since  $P_f(V)$  has only a single integral solution with  $x(s, t_i) = 1$  for each  $i = 1, 2, \dots, k$ , the result follows immediately.  $\square$

## 7 Applications in worst case analysis

In recent years there has been a lot of interest in the approximability of combinatorial optimization problems. Typically researchers propose a heuristic algorithm for an integer programming problem (a minimization problem) and compare the value of the heuristic to the value of the LP relaxation (or to the value of a dual feasible solution of the LP relaxation). A very nice and very general example of this approach is the  $2(1 - \frac{1}{|T|})$  approximation algorithm ( $T = \{v \in V : f(v) = 1\}$ ) proposed in Goemans and Williamson [4] for the problem  $IP_f(\emptyset)$  with  $f$  being proper (Conditions A) and taking values in  $\{0, 1\}$ . A corollary of their result is the bound  $\frac{IZ_f(\emptyset)}{Z_f(\emptyset)} \leq 2(1 - \frac{1}{|T|})$ . A distinct characteristic of their method is a reverse deletion step, in which edges that were added in the solution are deleted. Moreover, for the matching problem the bound is exact (the matching polyhedron  $P_f(V)$  is integral).

In this section we propose a new proof method that shows that  $\frac{IZ_f(\emptyset)}{Z_f(\emptyset)} \leq 2(1 - \frac{1}{|T|})$  for proper functions. The proof method gives rise to a new (and in our opinion more natural) algorithm that does not use reverse deletions and therefore it is easier to implement. Moreover, we remark that

the proof method is quite powerful as it can prove integrality of the matching polyhedron. It can also be used to prove the integrality of the multicut formulation for the minimum spanning tree problem and the branching polyhedron. Finally we use the parsimonious property to bound  $\frac{IZ_f(V)}{Z_f(V)}$  if  $c$  satisfies the triangle inequality.

### 7.1 A proof technique to bound the ratio $\frac{IZ_f(\emptyset)}{Z_f(\emptyset)}$

We consider problem  $IP_f(\emptyset)$  with  $f$  being a 0 – 1 proper function. Let  $T = \{v \in V : f(v) = 1\}$ . Our proof technique uses the crucial observation that a minimal solution to the problem must be a forest, and thus has at most  $|T| - 1$  edges.

**Theorem 14 (Goemans and Williamson [4])** *If  $f$  is a 0 – 1 proper function*

$$IZ_f(\emptyset) \leq 2\left(1 - \frac{1}{|T|}\right)Z_f(\emptyset).$$

**Proof :** For the purpose of contradiction we assume the contrary. Therefore, there exists a counter-example on the least number of nodes, with  $f$  proper and  $c$  integral. We may further assume that  $\sum_{e \in E} c(e)$  is minimal.

Suppose there is a  $v$  with  $f(v) = 0$ . Let  $f'$  denote the restriction of  $f$  on  $V \setminus \{v\}$ . It can easily be checked that  $f'$  is still proper. By the minimality of the counter-example,

$$IZ_{f'}(\emptyset) \leq 2\left(1 - \frac{1}{|T|}\right)Z_{f'}(\emptyset).$$

Since the optimal solution in  $IP_{f'}(\emptyset)$  is also feasible in  $IP_f(\emptyset)$ ,  $IZ_f(\emptyset) \leq IZ_{f'}(\emptyset)$ . From Theorem 4, by using the shortest path distances  $Z_f(\emptyset) = Z'_f(\emptyset)$  and  $Z_{f'}(\emptyset) = Z'_{f'}(\emptyset)$ . But,  $Z'_{f'}(\emptyset) = Z'_{f'}(\{v\})$ . By the parsimonious property  $Z'_f(\emptyset) = Z'_f(\{v\}) = Z'_{f'}(\emptyset)$ . Therefore,

$$IZ_f(\emptyset) \leq 2\left(1 - \frac{1}{|T|}\right)Z_f(\emptyset),$$

which is a contradiction. So we may assume  $f(v) = 1$  for all  $v$ .

If there is an edge  $e = (u, v) \in E$  with  $c_e = 0$ , then by contracting this edge, and treating  $\{u, v\}$  as a supernode, we restrict the problem to one of strictly smaller size. By the minimality of the counter-example, there exists a solution that satisfies the theorem. By introducing the edge  $(u, v)$ ,

with no extra cost since  $c_e = 0$ , if necessary, we obtain a solution feasible to the original problem and the theorem holds. Therefore, we may assume  $c_e > 0$  for all  $e$ .

Now let  $y(v) = \frac{1}{2}$  for all  $v$  and consider the cost function  $c'$  where  $c'_e = c_e - 1$  ( $c'_e \geq 0$  from the previous paragraph). By the minimality of  $c$ , there exist  $x, y'$  such that  $x \in IP_f(\emptyset)$ ,  $\sum_{S:e \in \delta(S)} y'(S) \leq c'_e$  and

$$\sum_e c'_e x_e \leq 2\left(1 - \frac{1}{|T|}\right) \sum_S y'(S) f(S).$$

Since  $f$  is 0-1,  $x_e$  corresponds to a forest and therefore,

$$\sum_e x_e \leq |T| - 1 = 2\left(1 - \frac{1}{|T|}\right) \sum_S y(v) = 2\left(1 - \frac{1}{|T|}\right) \sum_S f(v) y(v).$$

The last equality holds, since we have shown that we can assume  $f(v) = 1$ .

Let  $y^* = y' + y$ . Note that

$$\sum_{S:e \in \delta(S)} y^*(S) = \sum_{S:e \in \delta(S)} y(S) + \frac{1}{2} + \frac{1}{2} \leq c'_e + 1 = c_e$$

and so  $y^*$  is dual feasible. Therefore,

$$\sum_e c_e x_e = \sum_e c'_e x_e + \sum_e x_e \leq 2\left(1 - \frac{1}{|T|}\right) \sum_S y^*(S) f(S).$$

This is again a contradiction and the theorem follows.  $\square$

**Remarks:**

1. The dual variables  $y$  constructed in the proof are half-integral. We call a cut  $\delta(S)$  an  $f$ -cut if  $f(S) > 0$ . We can refine the previous theorem as follows. We have shown that  $IZ_f(\emptyset)$  is bounded above by  $2\left(1 - \frac{1}{|T|}\right)$  times the maximum half-integral  $c$ -packing of  $f$ -cuts. This observation has an interesting implication for the TSP. It is well-known that, if the cost function  $c$  satisfies the triangle inequality, the Christofides heuristic constructs a solution with objective value (denoted  $Z_C$ ) not more than  $3/2$  times of the optimum. This result has been strengthened further by Wolsey [17] and Shmoys and Williamson [15] who showed that  $Z_C \leq \left(\frac{3}{2} - \frac{1}{n}\right) Z_f(\emptyset)$ , where  $f$  corresponds to the TSP function. We can strengthen the inequality by replacing  $Z_f(\emptyset)$  with the value of the maximum half-integral  $c$ -packing (denoted by  $DZ_f(1/2)$ ). Note that all cuts are  $f$ -cuts in this instance and  $2 DZ_f(1/2) \leq DZ_f(\emptyset) =$

$Z_f(\emptyset)$ . From the above discussion, the solution to the minimum spanning tree is bounded above by  $2(1 - \frac{1}{|T|})DZ_f(1/2)$ . A well known result on matching (see [10]) says that the minimum matching on the set of odd nodes is bounded above by  $DZ_f(1/2)$ . Hence

$$Z_C \leq (3 - \frac{2}{|T|})DZ_f(1/2).$$

2. The proof only works for proper and not the more general parsimonious functions, because we want the parsimonious property to hold even if we create supernodes. Therefore, node subadditivity is not sufficient. Therefore, we need  $f$  to satisfy the full subadditivity and the QS property, which for the case of 0 – 1 functions is exactly the class of proper functions.
3. The above proof technique can be used to prove the integrality of the matching polyhedron, the multicut formulation for the minimum spanning tree problem and the branching polyhedron. For the matching polyhedron the difficult step is the case with  $c_e = 0$ , which can be handled using techniques from [10]. The final step is easy, since  $\sum_e x_e = \frac{|T|}{2} = \sum_v y(v)$ .

As an example of a different application of the proof method let us consider the multicut formulation of the MST. Let  $\Pi = \{S_1, \dots, S_{|\Pi|}\}$  be a partition of  $V$ .

$$\begin{aligned} IZ_{mcut} = \text{minimize} \quad & \sum_{e \in E} c_e x_e \\ \text{subject to} \quad & \sum_{e \in \delta(S_i, S_j); 1 \leq i < j \leq |\Pi|} x_e \geq |\Pi| - 1, \forall \Pi = \{S_1, \dots, S_{|\Pi|}\} \\ & x_e \in \{0, 1\}, \end{aligned}$$

Let  $Z_{mcut}$  be the LP relaxation and consider the dual problem.

$$\begin{aligned} Z_{mcut} = \text{maximize} \quad & \sum_{\Pi} (|\Pi| - 1)y(\Pi) \\ \text{subject to} \quad & \sum_{\Pi: e \in \delta(S_i, S_j)} y(\Pi) \leq c_e, \forall e \in E \\ & y(\Pi) \geq 0, \end{aligned}$$

We need to show that  $IZ_{mcut} = Z_{mcut}$ . We use an identical proof method, assuming the existence of a minimal counter-example. The only difference is that we update the dual variables as follows. Let  $\Pi = \{\{1\}, \{2\}, \dots, \{n\}\}$  and  $y(\Pi) = 1$ . Note that  $\sum_e x_e = n - 1 = \sum (|\Pi| - 1)y(\Pi)$ .

Although the proof method in Theorem 7.1 is non-algorithmic, it also leads to an algorithmic method to construct an approximate solution. It differs from the Goemans and Williamson's algorithm in that it needs a pre-processing step to compute pairwise shortest paths. With this in hand, we can discard all vertices with  $f(v) = 0$ . We call these vertices the Steiner vertices. This approach avoids the critical reverse deletion step in the Goemans and Williamson's algorithm, at the expense of computing pairwise shortest paths. Our algorithm is as follows:

**Approximation Algorithm for 0 – 1 proper functions**

1. Compute the pairwise shortest path distances for all pairs of non-Steiner nodes.
2. Discard the set of Steiner nodes. Select an edge with the least cost. Merge the two end nodes into a supernode, delete all edges joining these two nodes and update the costs of joining the supernodes.
3. Repeat step 2 until two supernodes merge to form a set  $S$  with  $f(S) = 0$ . Let  $e'$  be the last edge selected. If there are no more non-Steiner nodes, go to step 4. Else for all edges remaining, reduce the cost to  $c(e')$  and return to step 1.
4. Replace the edges selected in Step 2 and 3 by its corresponding shortest path in the original graph  $G$ .

For the Steiner tree problem, our algorithm emulates the MST heuristic on non-Steiner nodes with the pairwise shortest distance metric. In this respect Theorem 7.1 generalizes the well known fact that the MST heuristic gives a  $2(1 - \frac{1}{|T|})$  approximate solution to the minimum Steiner tree problem.

For arbitrary proper functions  $f$ , as Goemans and Williamson [4] observe, we can construct a feasible solution by utilizing Theorem 7.1. Let  $p_1 < p_2 < \dots < p_n$  be the distinct values of  $f$ , and for each  $i$ ,  $f_{p_i}(S) = 1$  if  $f(S) \geq p_i$  and 0 otherwise. Note that  $f_{p_i}$  is proper 0 – 1. By appending  $p_i - p_{i-1}$  ( $p_0 = 0$ ) copies of the approximate solution to  $f_{p_i}$ , for each  $i = 1, 2, \dots, n$ , we obtain a feasible solution which is within  $2\mathcal{H}(p_1, p_2, \dots, p_n)$  times of the optimal solution, where  $\mathcal{H}(p_1, p_2, \dots, p_n) = \sum_{i=1}^n \frac{p_i - p_{i-1}}{p_i}$ . Similarly for arbitrary QS functions we can use the results of [5]



to find

$$IZ_f(\emptyset) \leq 2\mathcal{H}(p_1, p_2, \dots, p_n)Z_f(\emptyset).$$

## 7.2 On the approximability of $IP_f(V)$

We next study the approximability of  $IP_f(V)$ . The recent research activity on approximation algorithms has so far concentrated on problem  $IP_f(\emptyset)$ , partly because it is difficult to construct a feasible integer solution to  $IP_f(V)$ . In fact, checking feasibility is usually NP-hard, as indicated for the case of the Hamiltonian-Cycle problem. Using our understanding of edge-splitting techniques and the parsimonious property, we can extend many of the approximation results to  $IZ_f(V)$ , when  $f$  is an even parsimonious function, and  $c$  satisfies the triangle inequality.

**Theorem 15** *If  $f$  is an even parsimonious function, and  $c$  satisfies triangle inequality, then*

$$IZ_f(V) \leq 2\mathcal{H}(p_1, p_2, \dots, p_n)Z_f(V).$$

**Proof :** Let  $f'$  be  $f/2$ . Then  $f'$  is again a parsimonious function. From [5] We first construct an approximate solution to  $IP_{f'}(\emptyset)$ , with  $x', y$  denoting the primal and dual solution respectively. Let  $x = 2x'$ . Then the graph corresponding to  $x$  is Eulerian, since each vertex has even degree. Note that

$$\begin{aligned} \mathcal{H}\left(\frac{p_1}{2}, \dots, \frac{p_n}{2}\right) &= \mathcal{H}(p_1, \dots, p_n), \text{ and} \\ \sum_e c_e x_e &= 2 \sum_e c_e x'_e \leq 4\mathcal{H}\left(\frac{p_1}{2}, \dots, \frac{p_n}{2}\right) \sum_S y(S) f'(S) = 2\mathcal{H}(p_1, \dots, p_n) \sum_S y(S) f(S). \end{aligned}$$

Since  $f$  is even and the graph corresponding to  $x$  is Eulerian, applying Corollary 1, we can use edge-splitting operations to construct a feasible solution to  $IZ_f(V)$ . By the triangle inequality, the cost of the constructed solution has not increased and therefore,

$$IZ_f(V) \leq 2\mathcal{H}(p_1, p_2, \dots, p_n)Z_f(\emptyset).$$

From the parsimonious property  $Z_f(\emptyset) = Z_f(V)$  and the theorem follows.  $\square$

**Remark:** For the case of the TSP, the previous theorem corresponds to the well-known fact that doubling the edges of the MST solution yields a 2-approximate solution to the TSP.

**Acknowledgement:**

We would like to thank Michel Goemans for pointing out his unpublished work in [6] to us.

## References

- [1] J. Edmonds, (1965). Maximum matching and a polyhedron with  $(0, 1)$ -vertices . *J. Res. Nat. Bur. Standards*, (69B), 125-130.
- [2] A. Frank, (1990). Packing paths, circuits and cuts-a survey, in *Paths, Flows and VLSI-Layout*. ed. B. Korte et. al., 47-100.
- [3] M.X. Goemans and D. Bertsimas, (1993). Survivable Networks, linear programming relaxations and the parsimonious property, *Mathematical Programming*, 60, 145-166.
- [4] M.X. Goemans and D.P. Williamson, (1992). A general approximation technique for constrained forest problems, *Proc. 3rd ACM-SIAM Symposium on Discrete Algorithms*. 307-316.
- [5] M.X. Goemans, A.V. Goldberg, S. Plotkin, D.B. Shmoys, E. Tardos and D.P. Williamson. (1994). Improved approximation algorithms for network design problems. *Proc. 5th ACM-SIAM Symposium of Discrete Algorithms*, to appear.
- [6] M.X. Goemans, A note on the parsimonious property, unpublished manuscript.
- [7] E.L. Johnson, (1975). On cut-set integer polyhedra, *Cahiers Centre Etudes Recherche Oper.* 17, 235-251.
- [8] L. Lovász, (1976). On some connectivity properties of Eulerian graphs. *Acta Math. Acad. Scient. Hung.*, 28, 129-138.
- [9] M.V. Lomonosov, (1979). Multiflow feasibility depending on cuts, *Graph Theory Newsletter*. 9, 4.
- [10] L. Lovász and M.D. Plummer, (1986). *Matching Theory*, North-Holland, Amsterdam.
- [11] R.Z. Norman, (1955). On the convex polyhedra of the symmetric travelling salesman problem. *Bull. Amer. Math. Soc.*, 61, 559.
- [12] B. Rothschild and A. Whinston, (1966). Feasibility of two-commodity network flows. *Operations Research*, 14, 1121-1129.

- [13] A. Schrijver, (1991). Short proofs on multicommodity flows and cuts, *J. Comb. Theory, Ser. B*, 53, 32-39.
- [14] P.D. Seymour, (1980). Four-terminal flows, *Networks*, 10, 79-86.
- [15] D. Shmoys and D.P. Williamson, (1990). Analyzing the Held-Karp TSP bound: a monotonicity property with application, *Information Processing Letters*, 35, 281-285.
- [16] D.P. Williamson, M.X. Goemans, M. Mihail and V. V. Vazirani, (1993). A primal-dual approximation algorithm for generalized Steiner network problems, *Proc. 25th Annual ACM Symposium on Theory of Computing*, 708-717.
- [17] L.A. Wolsey, (1980). Heuristic analysis, linear programming and Branch and bound, *Math. Prog. Study*, 13, 121-134.



MIT LIBRARIES



3 9080 00856632 2

2125 043





Date Due

APR 05 1987





